

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
14 June 2001 (14.06.2001)

PCT

(10) International Publication Number  
**WO 01/42435 A2**

- (51) International Patent Classification<sup>7</sup>: C12N 9/00
- (21) International Application Number: PCT/US00/33240
- (22) International Filing Date: 7 December 2000 (07.12.2000)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
60/169,428 7 December 1999 (07.12.1999) US
- (71) Applicant: **LEXICON GENETICS INCORPORATED** [US/US]; 4000 Research Forest Drive, The Woodlands, TX 77381 (US).
- (72) Inventors: **DONOHU, Gregory**; 41 Fallenstone Drive, The Woodlands, TX 77381 (US). **SCOVILLE, John**; 8222 Kingsbrook Dr. #543, Houston, TX 77024 (US). **TURNER, C., Alexander, Jr.**; 67 Winter Wheat Place, The Woodlands, TX 77381 (US). **FRIEDRICH, Glenn**; Breland & Breland, 2207 Hermann Drive, Houston, TX 77004 (US). **ZAMBROWICZ, Brian**; 18 Firethorne Place, The Woodlands, TX 77382 (US). **ABUIN, Alejandro**; 19 Belcarra Place, The Woodlands, TX 77382 (US). **SANDS, Arthur, T.**; 163 Bristol Bend Circle, The Woodlands, TX 77382 (US).
- (74) Agents: **ISHIMOTO, Lance, K. et al.**; Lexicon Genetics Incorporated, 4000 Research Forest Drive, The Woodlands, TX 77381 (US).
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- Published:**  
— Without international search report and to be republished upon receipt of that report.
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

WO 01/42435 A2

(54) Title: NOVEL HUMAN KINASE PROTEINS AND POLYNUCLEOTIDES ENCODING THE SAME

(57) Abstract: Novel human polynucleotide and polypeptide sequences are disclosed that can be used in therapeutic, diagnostic, and pharmacogenomic applications.

NOVEL HUMAN KINASE PROTEINS AND  
POLYNUCLEOTIDES ENCODING THE SAME

The present application claims the benefit of U.S.  
Provisional Application Number 60/169,428 which was filed on  
5 December 7, 1999 and is herein incorporated by reference in  
its entirety.

1. INTRODUCTION

The present invention relates to the discovery,  
identification, and characterization of novel human  
10 polynucleotides encoding proteins that share sequence  
similarity with animal kinases. The invention encompasses the  
described polynucleotides, host cell expression systems, the  
encoded proteins, fusion proteins, polypeptides and peptides,  
antibodies to the encoded proteins and peptides, and  
15 genetically engineered animals that either lack or over  
express the disclosed sequences, antagonists and agonists of  
the proteins, and other compounds that modulate the expression  
or activity of the proteins encoded by the disclosed genes  
that can be used for diagnosis, drug screening, clinical trial  
20 monitoring and the treatment of physiological disorders.

2. BACKGROUND OF THE INVENTION

Kinases mediate phosphorylation of a wide variety of  
proteins and compounds in the cell. Along with phosphatases,  
kinases are involved in a range of regulatory pathways. Given  
25 the physiological importance of kinases, they have been  
subject to intense scrutiny and are proven drug targets.

3. SUMMARY OF THE INVENTION

The present invention relates to the discovery,  
identification, and characterization of nucleotides that  
30 encode novel human proteins, and the corresponding amino acid  
sequences of these proteins. The novel human proteins (NHPs)  
described for the first time herein share structural  
similarity with animal kinases, including, but not limited to  
calcium/calmodulin-dependant protein kinases and

serine/threonine protein kinases. As such, the novel polynucleotides encode a novel kinase family having homologues and orthologs across a range of phyla and species.

The novel human polynucleotides described herein, encode  
5 open reading frames (ORFs) encoding proteins of 385 or 356 amino acids in length (see SEQ ID NOS: 2 and 4 respectively).

The invention also encompasses agonists and antagonists of the described NHPs, including small molecules, large molecules, mutant NHPs, or portions thereof that compete with  
10 native NHPs, NHP peptides, and antibodies, as well as nucleotide sequences that can be used to inhibit the expression of the described NHPs (e.g., antisense and ribozyme molecules, and gene or regulatory sequence replacement constructs) or to enhance the expression of the described NHP  
15 polynucleotides (e.g., expression constructs that place the described sequence under the control of a strong promoter system). The present invention also includes both transgenic animals that express a NHP transgene, and NHP "knock-outs" (which can be conditional) that do not express a functional  
20 NHP.

Also encompassed by the present invention are novel protein constructs engineered in such a way that they facilitate transport of the NHP to the target site, to the desired organ, across the cell membrane and/or to the nucleus where the NHP  
25 can exert its function activity. This goal may be achieved by coupling of the NHP to a cytokine or other ligand that would direct the NHP to the target organ and facilitate receptor mediated transport across the membrane into the cytosol.

Conjugation of NHPs to antibody molecules or their Fab  
30 fragments could be used to target cells bearing a particular epitope. Attaching the appropriate signal sequence to the NHP would also transport the NHP to the desired location within the cell. Alternatively targeting of NHP or its nucleic acid sequence might be achieved using liposome or lipid complex  
35 based delivery systems. Such technologies are described in Liposomes: A Practical Approach, R.R.C. New, Ed., Oxford University Press, New York and in U.S. Patents Nos. 4,594,595,

5,459,127, 5,948,767 and 6,110,490 and their respective disclosures which are herein incorporated by reference in their entirety.

Further, the present invention also relates to processes for identifying compounds that modulate, i.e., act as agonists or antagonists, of NHP expression and/or NHP product activity that utilize purified preparations of the described NHPs and/or NHP product, or cells expressing the same. Such compounds can be used as therapeutic agents for the treatment of any of a wide variety of symptoms associated with biological disorders or imbalances.

#### 4. DESCRIPTION OF THE SEQUENCE LISTING AND FIGURES

The Sequence Listing provides the sequence of the novel human ORFs encoding the described novel human kinase proteins. SEQ ID NO:5 describes a full length ORF and flanking regions.

#### 5. DETAILED DESCRIPTION OF THE INVENTION

The NHPs, described for the first time herein, are novel protein that is expressed in, *inter alia*, human cell lines, and human fetal brain, pituitary, cerebellum, thymus, spleen, trachea, lung, testis, thyroid, adrenal gland, pancreas, colon, uterus, and fetal kidney cells. The described sequences were compiled from gene trapped cDNAs, ESTs, a and human thymus cDNA library, (Edge Biosystems, Gaithersburg, MD).

The present invention encompasses the nucleotides presented in the Sequence Listing, host cells expressing such nucleotides, the expression products of such nucleotides, and: (a) nucleotides that encode mammalian homologs of the described sequences, including the specifically described NHPs, and the NHP products; (b) nucleotides that encode one or more portions of an NHP that correspond to functional domains, and the polypeptide products specified by such nucleotide sequences, including but not limited to the novel regions of any active domain(s); (c) isolated nucleotides that encode mutant versions, engineered or naturally occurring, of the

described NHPs in which all or a part of at least one domain is deleted or altered, and the polypeptide products specified by such nucleotide sequences, including but not limited to soluble proteins and peptides in which all or a portion of the signal sequence is deleted; (d) nucleotides that encode chimeric fusion proteins containing all or a portion of a coding region of a NHP, or one of its domains (e.g., a receptor/ligand binding domain, accessory protein/self-association domain, etc.) fused to another peptide or polypeptide; or (e) therapeutic or diagnostic derivatives of the described polynucleotides such as oligonucleotides, antisense polynucleotides, ribozymes, dsRNA, or gene therapy constructs comprising a sequence first disclosed in the Sequence Listing. As discussed above, the present invention includes: (a) the human DNA sequences presented in the Sequence Listing (and vectors comprising the same) and additionally contemplates any nucleotide sequence encoding a contiguous NHP open reading frame (ORF) that hybridizes to a complement of a DNA sequence presented in the Sequence Listing under highly stringent conditions, e.g., hybridization to filter-bound DNA in 0.5 M NaHPO<sub>4</sub>, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1xSSC/0.1% SDS at 68°C (Ausubel F.M. et al., eds., 1989, Current Protocols in Molecular Biology, Vol. I, Green Publishing Associates, Inc., and John Wiley & sons, Inc., New York, at p. 2.10.3) and encodes a functionally equivalent gene product. Additionally contemplated are any nucleotide sequences that hybridize to the complement of the DNA sequence that encode and express an amino acid sequence presented in the Sequence Listing under moderately stringent conditions, e.g., washing in 0.2xSSC/0.1% SDS at 42°C (Ausubel et al., 1989, *supra*), yet still encode a functionally equivalent NHP product. Functional equivalents of a NHP include naturally occurring NHPs present in other species and mutant NHPs whether naturally occurring or engineered (by site directed mutagenesis, gene shuffling, directed evolution as described in, for example, U.S. Patent No. 5,837,458). The invention also includes degenerate

nucleic acid variants of the disclosed NHP polynucleotide sequences.

Additionally contemplated are polynucleotides encoding NHP ORFs, or their functional equivalents, encoded by  
5 polynucleotide sequences that are about 99, 95, 90, or about 85 percent similar to corresponding regions of SEQ ID NO:1 (as measured by BLAST sequence comparison analysis using, for example, the GCG sequence analysis package using default parameters).

10 The invention also includes nucleic acid molecules, preferably DNA molecules, that hybridize to, and are therefore the complements of, the described NHP encoding polynucleotides. Such hybridization conditions can be highly stringent or less highly stringent, as described above. In  
15 instances where the nucleic acid molecules are deoxyoligonucleotides ("DNA oligos"), such molecules are generally about 16 to about 100 bases long, or about 20 to about 80, or about 34 to about 45 bases long, or any variation or combination of sizes represented therein that incorporate a  
20 contiguous region of sequence first disclosed in the Sequence Listing. Such oligonucleotides can be used in conjunction with the polymerase chain reaction (PCR) to screen libraries, isolate clones, and prepare cloning and sequencing templates, etc.

25 Alternatively, such NHP oligonucleotides can be used as hybridization probes for screening libraries, and assessing gene expression patterns (particularly using a micro array or high-throughput "chip" format). Additionally, a series of the described NHP oligonucleotide sequences, or the complements  
30 thereof, can be used to represent all or a portion of the described NHP sequences. The oligonucleotides, typically between about 16 to about 40 (or any whole number within the stated range) nucleotides in length may partially overlap each other and/or the NHP sequence may be represented using  
35 oligonucleotides that do not overlap. Accordingly, the described NHP polynucleotide sequences shall typically comprise at least about two or three distinct oligonucleotide

sequences of at least about 18, and preferably about 25, nucleotides in length that are each first disclosed in the described Sequence Listing. Such oligonucleotide sequences may begin at any nucleotide present within a sequence in the Sequence Listing and proceed in either a sense (5'-to-3') orientation vis-a-vis the described sequence or in an antisense orientation.

For oligonucleotide probes, highly stringent conditions may refer, e.g., to washing in 6xSSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligos), 48°C (for 17-base oligos), 55°C (for 20-base oligos), and 60°C (for 23-base oligos). These nucleic acid molecules may encode or act as NHP sequence antisense molecules, useful, for example, in NHP gene regulation (for and/or as antisense primers in amplification reactions of NHP nucleic acid sequences). With respect to NHP gene regulation, such techniques can be used to regulate biological functions. Further, such sequences can be used as part of ribozyme and/or triple helix sequences that are also useful for NHP gene regulation.

Inhibitory antisense or double stranded oligonucleotides can additionally comprise at least one modified base moiety which is selected from the group including but not limited to 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-

5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

The antisense oligonucleotide can also comprise at least one modified sugar moiety selected from the group including  
5 but not limited to arabinose, 2-fluoroarabinose, xylulose, and hexose.

In yet another embodiment, the antisense oligonucleotide will comprise at least one modified phosphate backbone selected from the group consisting of a phosphorothioate, a  
10 phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

In yet another embodiment, the antisense oligonucleotide is an  $\alpha$ -anomeric oligonucleotide. An  $\alpha$ -anomeric  
15 oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual  $\beta$ -units, the strands run parallel to each other (Gautier et al., 1987, Nucl. Acids Res. 15:6625-6641). The oligonucleotide is a 2'-O-methylribonucleotide (Inoue et al., 1987, Nucl. Acids Res.  
20 15:6131-6148), or a chimeric RNA-DNA analogue (Inoue et al., 1987, FEBS Lett. 215:327-330). Alternatively, double stranded RNA can be used to disrupt the expression and function of a targeted NHP.

Oligonucleotides of the invention can be synthesized by  
25 standard methods known in the art, e.g. by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides can be synthesized by the method of Stein et al. (1988, Nucl. Acids Res. 16:3209), and  
30 methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin et al., 1988, Proc. Natl. Acad. Sci. U.S.A. 85:7448-7451), etc.

Low stringency conditions are well known to those of skill in the art, and will vary predictably depending on the  
35 specific organisms from which the library and the labeled sequences are derived. For guidance regarding such conditions see, for example, Sambrook et al., 1989, Molecular Cloning, A



Laboratory Manual (and periodic updates thereof), Cold Springs Harbor Press, N.Y.; and Ausubel *et al.*, 1989, Current Protocols in Molecular Biology, Green Publishing Associates and Wiley Interscience, N.Y.

- 5        Alternatively, suitably labeled NHP nucleotide probes can be used to screen a human genomic library using appropriately stringent conditions or by PCR. The identification and characterization of human genomic clones is helpful for identifying polymorphisms (including, but not limited to, nucleotide repeats, microsatellite alleles, single nucleotide polymorphisms, or coding single nucleotide polymorphisms), determining the genomic structure of a given locus/allele, and designing diagnostic tests. For example, sequences derived from regions adjacent to the intron/exon boundaries of the human gene can be used to design primers for use in amplification assays to detect mutations within the exons, introns, splice sites (e.g., splice acceptor and/or donor sites), etc., that can be used in diagnostics and pharmacogenomics.
- 20        Further, a NHP sequence homolog can be isolated from nucleic acid from an organism of interest by performing PCR using two degenerate or "wobble" oligonucleotide primer pools designed on the basis of amino acid sequences within the NHP products disclosed herein. The template for the reaction may be total RNA, mRNA, and/or cDNA obtained by reverse transcription of mRNA prepared from, for example, human or non-human cell lines or tissue known or suspected to express an allele of a NHP gene.

- 30        The PCR product can be subcloned and sequenced to ensure that the amplified sequences represent the sequence of the desired NHP gene. The PCR fragment can then be used to isolate a full length cDNA clone by a variety of methods. For example, the amplified fragment can be labeled and used to screen a cDNA library, such as a bacteriophage cDNA library.
- 35        Alternatively, the labeled fragment can be used to isolate genomic clones via the screening of a genomic library.

PCR technology can also be used to isolate full length cDNA sequences. For example, RNA can be isolated, following standard procedures, from an appropriate cellular or tissue source (*i.e.*, one known, or suspected, to express a NHP gene).

5 A reverse transcription (RT) reaction can be performed on the RNA using an oligonucleotide primer specific for the most 5' end of the amplified fragment for the priming of first strand synthesis. The resulting RNA/DNA hybrid may then be "tailed" using a standard terminal transferase reaction, the hybrid may  
10 be digested with RNase H, and second strand synthesis may then be primed with a complementary primer. Thus, cDNA sequences upstream of the amplified fragment can be isolated. For a review of cloning strategies that can be used, see *e.g.*, Sambrook *et al.*, 1989, *supra*.

15 A cDNA encoding a mutant NHP gene can be isolated, for example, by using PCR. In this case, the first cDNA strand may be synthesized by hybridizing an oligo-dT oligonucleotide to mRNA isolated from tissue known or suspected to be expressed in an individual putatively carrying a mutant NHP  
20 allele, and by extending the new strand with reverse transcriptase. The second strand of the cDNA is then synthesized using an oligonucleotide that hybridizes specifically to the 5' end of the normal gene. Using these two primers, the product is then amplified via PCR, optionally  
25 cloned into a suitable vector, and subjected to DNA sequence analysis through methods well known to those of skill in the art. By comparing the DNA sequence of the mutant NHP allele to that of a corresponding normal NHP allele, the mutation(s) responsible for the loss or alteration of function of the  
30 mutant NHP gene product can be ascertained.

Alternatively, a genomic library can be constructed using DNA obtained from an individual suspected of or known to carry a mutant NHP allele (*e.g.*, a person manifesting a NHP-associated phenotype such as, for example, immune disorders,  
35 obesity, high blood pressure, etc.), or a cDNA library can be constructed using RNA from a tissue known, or suspected, to express a mutant NHP allele. A normal NHP sequence, or any

suitable fragment thereof, can then be labeled and used as a probe to identify the corresponding mutant NHP allele in such libraries. Clones containing mutant NHP sequences can then be purified and subjected to sequence analysis according to methods well known to those skilled in the art.

Additionally, an expression library can be constructed utilizing cDNA synthesized from, for example, RNA isolated from a tissue known, or suspected, to express a mutant NHP allele in an individual suspected of or known to carry such a mutant allele. In this manner, gene products made by the putatively mutant tissue may be expressed and screened using standard antibody screening techniques in conjunction with antibodies raised against a normal NHP product, as described below. (For screening techniques, see, for example, Harlow, E. and Lane, eds., 1988, "Antibodies: A Laboratory Manual", Cold Spring Harbor Press, Cold Spring Harbor.) Additionally, screening can be accomplished by screening with labeled NHP fusion proteins, such as, for example, alkaline phosphatase-NHP or NHP-alkaline phosphatase fusion proteins. In cases where a NHP mutation results in an expressed gene product with altered function (e.g., as a result of a missense or a frameshift mutation), polyclonal antibodies to a NHP are likely to cross-react with a corresponding mutant NHP gene product. Library clones detected via their reaction with such labeled antibodies can be purified and subjected to sequence analysis according to methods well known in the art.

An additional application of the described novel human polynucleotide sequences is their use in the molecular mutagenesis/evolution of proteins that are at least partially encoded by the described novel sequences using, for example, polynucleotide shuffling or related methodologies. Such approaches are described in U.S. Patents Nos. 5,830,721 and 5,837,458 which are herein incorporated by reference in their entirety.

The invention also encompasses (a) DNA vectors that contain any of the foregoing NHP coding sequences and/or their complements (i.e., antisense); (b) DNA expression vectors that

contain any of the foregoing NHP coding sequences operatively associated with a regulatory element that directs the expression of the coding sequences (for example, baculo virus as described in U.S. Patent No. 5,869,336 herein incorporated by reference); (c) genetically engineered host cells that contain any of the foregoing NHP coding sequences operatively associated with a regulatory element that directs the expression of the coding sequences in the host cell; and (d) genetically engineered host cells that express an endogenous NHP gene under the control of an exogenously introduced regulatory element (i.e., gene activation). As used herein, regulatory elements include, but are not limited to, inducible and non-inducible promoters, enhancers, operators and other elements known to those skilled in the art that drive and regulate expression. Such regulatory elements include but are not limited to the human cytomegalovirus (hCMV) immediate early gene, regulatable, viral elements (particularly retroviral LTR promoters), the early or late promoters of SV40 adenovirus, the *lac* system, the *trp* system, the *TAC* system, the *TRC* system, the major operator and promoter regions of phage lambda, the control regions of fd coat protein, the promoter for 3-phosphoglycerate kinase (PGK), the promoters of acid phosphatase, and the promoters of the yeast  $\alpha$ -mating factors.

Where, as in the present instance, some of the described NHP peptides or polypeptides are thought to be cytoplasmic proteins, expression systems can be engineered that produce soluble derivatives of a NHP (corresponding to a NHP extracellular and/or intracellular domains, or truncated polypeptides lacking one or more hydrophobic domains) and/or NHP fusion protein products (especially NHP-Ig fusion proteins, i.e., fusions of a NHP domain to an IgFc), NHP antibodies, and anti-idiotypic antibodies (including Fab fragments) that can be used in therapeutic applications. Preferably, the above expression systems are engineered to allow the desired peptide or polypeptide to be recovered from the culture media.

The present invention also encompasses antibodies and anti-idiotypic antibodies (including Fab fragments), antagonists and agonists of a NHP, as well as compounds or nucleotide constructs that inhibit expression of a NHP  
5 sequence (transcription factor inhibitors, antisense and ribozyme molecules, or gene or regulatory sequence replacement constructs), or promote the expression of a NHP (e.g., expression constructs in which NHP coding sequences are operatively associated with expression control elements such  
10 as promoters, promoter/enhancers, etc.).

The NHPs or NHP peptides, NHP fusion proteins, NHP nucleotide sequences, antibodies, antagonists and agonists can be useful for the detection of mutant NHPs or inappropriately expressed NHPs for the diagnosis of disease. The NHP proteins  
15 or peptides, NHP fusion proteins, NHP nucleotide sequences, host cell expression systems, antibodies, antagonists, agonists and genetically engineered cells and animals can be used for screening for drugs (or high throughput screening of combinatorial libraries) effective in the treatment of the  
20 symptomatic or phenotypic manifestations of perturbing the normal function of a NHP in the body. The use of engineered host cells and/or animals can offer an advantage in that such systems allow not only for the identification of compounds that bind to the endogenous receptor/ligand of a NHP, but can  
25 also identify compounds that trigger NHP-mediated activities or pathways.

Finally, the NHP products can be used as therapeutics. For example, soluble derivatives such as NHP peptides/domains corresponding the NHPs, NHP fusion protein products  
30 (especially NHP-Ig fusion proteins, i.e., fusions of a NHP, or a domain of a NHP, to an IgFc), NHP antibodies and anti-idiotypic antibodies (including Fab fragments), antagonists or agonists (including compounds that modulate or act on downstream targets in a NHP-mediated pathway) can be used to  
35 directly treat diseases or disorders. For instance, the administration of an effective amount of soluble NHP, or a NHP-IgFc fusion protein or an anti-idiotypic antibody (or its

Fab) that mimics the NHP could activate or effectively antagonize the endogenous NHP or a protein interactive therewith. Nucleotide constructs encoding such NHP products can be used to genetically engineer host cells to express such products *in vivo*; these genetically engineered cells function as "bioreactors" in the body delivering a continuous supply of a NHP, a NHP peptide, or a NHP fusion protein to the body. Nucleotide constructs encoding functional NHPs, mutant NHPs, as well as antisense and ribozyme molecules can also be used in "gene therapy" approaches for the modulation of NHP expression. Thus, the invention also encompasses pharmaceutical formulations and methods for treating biological disorders.

Various aspects of the invention are described in greater detail in the subsections below.

#### 5.1 THE NHP SEQUENCES

The cDNA sequences and corresponding deduced amino acid sequences of the described NHPs are presented in the Sequence Listing. The NHP nucleotide sequences were obtained from a human cDNA library using probes and/or primers generated from human gene trapped sequence tags.

Expression analysis has provided evidence that the described NHPs can be expressed in human tissues as well as gene trapped human cells. In addition to the serine/threonine kinases, the described NHPs also share significant similarity to a range of additional kinase families such as calcium dependent protein kinases from a range of phyla and species. Given the physiological importance of protein kinases, they have been subject to intense scrutiny as exemplified and discussed in U.S. Patent No. 5,817,479 herein incorporated by reference in its entirety. Additionally, an oligonucleotide or polynucleotide sequence first disclosed in at least a portion of one or more of the sequences of SEQ ID NOS: 1-5 can be used as a hybridization probe in conjunction with a solid support matrix/substrate (resins, beads, membranes, plastics, polymers, metal or metallized substrates, crystalline or

polycrystalline substrates, etc.). Of particular note are spatially addressable arrays (i.e., gene chips, microtiter plates, etc.) of oligonucleotides and polynucleotides, or corresponding oligopeptides and polypeptides, wherein at least one of the biopolymers present on the spatially addressable array comprises an oligonucleotide or polynucleotide sequence first disclosed in at least one of the sequences of SEQ ID NOS: 1-5, or an amino acid sequence encoded thereby. Methods for attaching biopolymers to, or synthesizing biopolymers on, solid support matrices, and conducting binding studies thereon are disclosed in, *inter alia*, U.S. Patent Nos. 5,700,637, 5,556,752, 5,744,305, 4,631,211, 5,445,934, 5,252,743, 4,713,326, 5,424,186, and 4,689,405 the disclosures of which are herein incorporated by reference in their entirety.

Addressable arrays comprising sequences first disclosed in SEQ ID NOS:1-5 can be used to identify and characterize the temporal and tissue specific expression of a gene. These addressable arrays incorporate oligonucleotide sequences of sufficient length to confer the required specificity, yet be within the limitations of the production technology. The length of these probes is within a range of between about 8 to about 2000 nucleotides. Preferably the probes consist of 60 nucleotides and more preferably 25 nucleotides from the sequences first disclosed in SEQ ID NOS:1-5.

For example, a series of the described oligonucleotide sequences, or the complements thereof, can be used in chip format to represent all or a portion of the described sequences. The oligonucleotides, typically between about 16 to about 40 (or any whole number within the stated range) nucleotides in length can partially overlap each other and/or the sequence may be represented using oligonucleotides that do not overlap. Accordingly, the described GTS polynucleotide sequences shall typically comprise at least about two or three distinct oligonucleotide sequences of at least about 8 nucleotides in length that are each first disclosed in the described Sequence Listing. Such oligonucleotide sequences can begin at any nucleotide present within a sequence in the

Sequence Listing and proceed in either a sense (5'-to-3') orientation vis-a-vis the described sequence or in an antisense orientation.

Microarray-based analysis allows the discovery of broad  
5 patterns of genetic activity, providing new understanding of  
gene functions and generating novel and unexpected insight  
into transcriptional processes and biological mechanisms. The  
use of addressable arrays comprising sequences first disclosed  
in SEQ ID NOS:1-5 provides detailed information about  
10 transcriptional changes involved in a specific pathway,  
potentially leading to the identification of novel components  
or gene functions that manifest themselves as novel  
phenotypes.

Probes consisting of sequences first disclosed in SEQ ID  
15 NOS:1-5 can also be used in the identification, selection and  
validation of novel molecular targets for drug discovery. The  
use of these unique sequences permits the direct confirmation  
of drug targets and recognition of drug dependent changes in  
gene expression that are modulated through pathways distinct  
20 from the drugs intended target. These unique sequences  
therefore also have utility in defining and monitoring both  
drug action and toxicity.

As an example of utility, the sequences first disclosed  
in SEQ ID NOS:1-5 can be utilized in microarrays or other  
25 assay formats, to screen collections of genetic material from  
patients who have a particular medical condition. These  
investigations can also be carried out using the sequences  
first disclosed in SEQ ID NOS:1-5 *in silico* and by comparing  
previously collected genetic databases and the disclosed  
30 sequences using computer software known to those in the art.

Thus the sequences first disclosed in SEQ ID NOS:1-5 can  
be used to identify mutations associated with a particular  
disease and also as a diagnostic or prognostic assay.

Although the presently described sequences have been  
35 specifically described using nucleotide sequence, it should be  
appreciated that each of the sequences can uniquely be  
described using any of a wide variety of additional structural



attributes, or combinations thereof. For example, a given sequence can be described by the net composition of the nucleotides present within a given region of the sequence in conjunction with the presence of one or more specific  
5 oligonucleotide sequence(s) first disclosed in the SEQ ID NOS: 1-5. Alternatively, a restriction map specifying the relative positions of restriction endonuclease digestion sites, or various palindromic or other specific oligonucleotide sequences can be used to structurally describe a given  
10 sequence. Such restriction maps, which are typically generated by widely available computer programs (e.g., the University of Wisconsin GCG sequence analysis package, SEQUENCHER 3.0, Gene Codes Corp., Ann Arbor, MI, etc.), can optionally be used in conjunction with one or more discrete  
15 nucleotide sequence(s) present in the sequence that can be described by the relative position of the sequence relative to one or more additional sequence(s) or one or more restriction sites present in the disclosed sequence.

## 20 5.2 NHPS AND NHP POLYPEPTIDES

NHPS, polypeptides, peptide fragments, mutated, truncated, or deleted forms of the NHPS, and/or NHP fusion proteins can be prepared for a variety of uses. These uses include, but are not limited to, the generation of antibodies,  
25 as reagents in diagnostic assays, for the identification of other cellular gene products related to a NHP, as reagents in assays for screening for compounds that can be as pharmaceutical reagents useful in the therapeutic treatment of mental, biological, or medical disorders and disease.

30 The Sequence Listing discloses the amino acid sequence encoded by the described NHP-encoding polynucleotides. The NHP has an initiator methionine in a DNA sequence context consistent with eucaryotic translation initiation site.

The NHP amino acid sequences of the invention include the  
35 amino acid sequences presented in the Sequence Listing as well as analogues and derivatives thereof. Further, corresponding NHP homologues from other species are encompassed by the

invention. In fact, any NHP protein encoded by the NHP nucleotide sequences described above are within the scope of the invention, as are any novel polynucleotide sequences encoding all or any novel portion of an amino acid sequence presented in the Sequence Listing. The degenerate nature of the genetic code is well known, and, accordingly, each amino acid presented in the Sequence Listing, is generically representative of the well known nucleic acid "triplet" codon, or in many cases codons, that can encode the amino acid. As such, as contemplated herein, the amino acid sequences presented in the Sequence Listing, when taken together with the genetic code (see, for example, Table 4-1 at page 109 of "Molecular Cell Biology", 1986, J. Darnell et al. eds., Scientific American Books, New York, NY, herein incorporated by reference) are generically representative of all the various permutations and combinations of nucleic acid sequences that can encode such amino acid sequences.

The invention also encompasses proteins that are functionally equivalent to the NHPs encoded by the presently described nucleotide sequences as judged by any of a number of criteria, including, but not limited to, the ability to bind and modify a NHP substrate, or the ability to effect an identical or complementary downstream pathway, or a change in cellular metabolism (e.g., proteolytic activity, ion flux, tyrosine phosphorylation, etc.). Such functionally equivalent NHP proteins include, but are not limited to, additions or substitutions of amino acid residues within the amino acid sequence encoded by a NHP nucleotide sequence described above, but which result in a silent change, thus producing a functionally equivalent gene product. Amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine,

and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid.

A variety of host-expression vector systems can be used  
5 to express the NHP nucleotide sequences of the invention. Where the NHP peptide or polypeptide can exist, or has been engineered to exist, as a soluble or secreted molecule, the soluble NHP peptide or polypeptide can be recovered from the culture media. Such expression systems also encompass  
10 engineered host cells that express a NHP, or functional equivalent, *in situ*. Purification or enrichment of a NHP from such expression systems can be accomplished using appropriate detergents and lipid micelles and methods well known to those skilled in the art. However, such engineered host cells  
15 themselves may be used in situations where it is important not only to retain the structural and functional characteristics of the NHP, but to assess biological activity, *e.g.*, in drug screening assays.

The expression systems that may be used for purposes of  
20 the invention include but are not limited to microorganisms such as bacteria (*e.g.*, *E. coli*, *B. subtilis*) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing NHP nucleotide sequences; yeast (*e.g.*, *Saccharomyces*, *Pichia*) transformed with recombinant  
25 yeast expression vectors containing NHP nucleotide sequences; insect cell systems infected with recombinant virus expression vectors (*e.g.*, baculovirus) containing NHP sequences; plant cell systems infected with recombinant virus expression  
30 vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (*e.g.*, Ti plasmid) containing NHP nucleotide sequences; or mammalian cell systems (*e.g.*, COS, CHO, BHK, 293, 3T3) harboring recombinant expression constructs containing promoters derived from the genome of mammalian  
35 cells (*e.g.*, metallothionein promoter) or from mammalian viruses (*e.g.*, the adenovirus late promoter; the vaccinia virus 7.5K promoter).

In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the NHP product being expressed. For example, when a large quantity of such a protein is to be produced for the

5 generation of pharmaceutical compositions of or containing NHP, or for raising antibodies to a NHP, vectors that direct the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include, but are not limited, to the *E. coli* expression vector pUR278

10 (Ruther et al., 1983, EMBO J. 2:1791), in which a NHP coding sequence may be ligated individually into the vector in frame with the *lacZ* coding region so that a fusion protein is produced; pIN vectors (Inouye & Inouye, 1985, Nucleic Acids Res. 13:3101-3109; Van Heeke & Schuster, 1989, J. Biol. Chem.

15 264:5503-5509); and the like. pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in

20 the presence of free glutathione. The PGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned target gene product can be released from the GST moiety.

In an insect system, *Autographa californica* nuclear

25 polyhydrosis virus (AcNPV) is used as a vector to express foreign genes. The virus grows in *Spodoptera frugiperda* cells. A NHP encoding polynucleotide sequence can be cloned individually into non-essential regions (for example the polyhedrin gene) of the virus and placed under control of an

30 AcNPV promoter (for example the polyhedrin promoter). Successful insertion of NHP gene coding sequence will result in inactivation of the polyhedrin gene and production of non-occluded recombinant virus (i.e., virus lacking the proteinaceous coat coded for by the polyhedrin gene). These

35 recombinant viruses are then used to infect *Spodoptera frugiperda* cells in which the inserted gene is expressed

(e.g., see Smith et al., 1983, J. Virol. 46: 584; Smith, U.S. Patent No. 4,215,051).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an  
5 adenovirus is used as an expression vector, the NHP nucleotide sequence of interest may be ligated to an adenovirus transcription/translation control complex, e.g., the late promoter and tripartite leader sequence. This chimeric gene can then be inserted in the adenovirus genome by *in vitro* or  
10 *in vivo* recombination. Insertion in a non-essential region of the viral genome (e.g., region E1 or E3) will result in a recombinant virus that is viable and capable of expressing a NHP product in infected hosts (e.g., See Logan & Shenk, 1984, Proc. Natl. Acad. Sci. USA 81:3655-3659). Specific initiation  
15 signals may also be required for efficient translation of inserted NHP nucleotide sequences. These signals include the ATG initiation codon and adjacent sequences. In cases where an entire NHP gene or cDNA, including its own initiation codon and adjacent sequences, is inserted into the appropriate  
20 expression vector, no additional translational control signals may be needed. However, in cases where only a portion of a NHP coding sequence is inserted, exogenous translational control signals, including, perhaps, the ATG initiation codon, must be provided. Furthermore, the initiation codon must be  
25 in phase with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of  
30 appropriate transcription enhancer elements, transcription terminators, etc. (See Bittner et al., 1987, Methods in Enzymol. 153:516-544).

In addition, a host cell strain may be chosen that modulates the expression of the inserted sequences, or  
35 modifies and processes the gene product in the specific fashion desired. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may be

important for the function of the protein. Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of proteins and gene products. Appropriate cell lines or host systems can be  
5 chosen to ensure the correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells which possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used. Such  
10 mammalian host cells include, but are not limited to, CHO, VERO, BHK, HeLa, COS, MDCK, 293, 3T3, WI38, and in particular, human cell lines.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell  
15 lines that stably express the NHP sequences described above can be engineered. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with DNA controlled by appropriate expression control elements (e.g., promoter, enhancer sequences,  
20 transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the introduction of the foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers  
25 resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines. This method may advantageously be used to engineer cell lines which express the NHP product. Such engineered cell lines may  
30 be particularly useful in screening and evaluation of compounds that affect the endogenous activity of the NHP product.

A number of selection systems can be used, including but not limited to the herpes simplex virus thymidine kinase  
35 (Wigler, et al., 1977, Cell 11:223), hypoxanthine-guanine phosphoribosyltransferase (Szybalska & Szybalski, 1962, Proc. Natl. Acad. Sci. USA 48:2026), and adenine

phosphoribosyltransferase (Lowy, et al., 1980, Cell 22:817) genes can be employed in tk<sup>-</sup>, hgp<sup>-</sup> or apr<sup>-</sup> cells, respectively. Also, antimetabolite resistance can be used as the basis of selection for the following genes: dhfr, which  
5 confers resistance to methotrexate (Wigler, et al., 1980, Natl. Acad. Sci. USA 77:3567; O'Hare, et al., 1981, Proc. Natl. Acad. Sci. USA 78:1527); gpt, which confers resistance to mycophenolic acid (Mulligan & Berg, 1981, Proc. Natl. Acad. Sci. USA 78:2072); neo, which confers resistance to the  
10 aminoglycoside G-418 (Colberre-Garapin, et al., 1981, J. Mol. Biol. 150:1); and hyg<sup>r</sup>, which confers resistance to hygromycin (Santerre, et al., 1984, Gene 30:147).

Alternatively, any fusion protein can be readily purified by utilizing an antibody specific for the fusion protein being  
15 expressed. For example, a system described by Janknecht et al. allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (Janknecht, et al., 1991, Proc. Natl. Acad. Sci. USA 88:8972-8976). In this system, the gene of interest is subcloned into a vaccinia  
20 recombination plasmid such that the gene's open reading frame is translationally fused to an amino-terminal tag consisting of six histidine residues. Extracts from cells infected with recombinant vaccinia virus are loaded onto Ni<sup>2+</sup> nitriloacetic acid-agarose columns and histidine-tagged proteins are  
25 selectively eluted with imidazole-containing buffers.

### 5.3 ANTIBODIES TO NHP PRODUCTS

Antibodies that specifically recognize one or more epitopes of a NHP, or epitopes of conserved variants of a NHP,  
30 or peptide fragments of a NHP are also encompassed by the invention. Such antibodies include but are not limited to polyclonal antibodies, monoclonal antibodies (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab')<sub>2</sub> fragments, fragments produced by a Fab  
35 expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above.

The antibodies of the invention can be used, for example, in the detection of NHP in a biological sample and may, therefore, be utilized as part of a diagnostic or prognostic technique whereby patients may be tested for abnormal amounts of NHP. Such antibodies may also be utilized in conjunction with, for example, compound screening schemes for the evaluation of the effect of test compounds on expression and/or activity of a NHP gene product. Additionally, such antibodies can be used in conjunction gene therapy to, for example, evaluate the normal and/or engineered NHP-expressing cells prior to their introduction into the patient. Such antibodies may additionally be used as a method for the inhibition of abnormal NHP activity. Thus, such antibodies may, therefore, be utilized as part of treatment methods.

For the production of antibodies, various host animals may be immunized by injection with the NHP, a NHP peptide (e.g., one corresponding the a functional domain of a NHP), truncated NHP polypeptides (NHP in which one or more domains have been deleted), functional equivalents of the NHP or mutated variant of the NHP. Such host animals may include but are not limited to pigs, rabbits, mice, goats, and rats, to name but a few. Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to Freund's adjuvant (complete and incomplete), mineral salts such as aluminum hydroxide or aluminum phosphate, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and *Corynebacterium parvum*.

Alternatively, the immune response could be enhanced by combination and or coupling with molecules such as keyhole limpet hemocyanin, tetanus toxoid, diphtheria toxoid, ovalbumin, cholera toxin or fragments thereof. Polyclonal antibodies are heterogeneous populations of antibody molecules derived from the sera of the immunized animals.

Monoclonal antibodies, which are homogeneous populations of antibodies to a particular antigen, can be obtained by any



technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique of Kohler and Milstein, (1975, *Nature* 256:495-497; and U.S. Patent No.

5 4,376,110), the human B-cell hybridoma technique (Kosbor et al., 1983, *Immunology Today* 4:72; Cole et al., 1983, *Proc. Natl. Acad. Sci. USA* 80:2026-2030), and the EBV-hybridoma technique (Cole et al., 1985, *Monoclonal Antibodies And Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96). Such antibodies may  
10 be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. The hybridoma producing the mAb of this invention may be cultivated in vitro or in vivo. Production of high titers of mAbs in vivo makes this the presently preferred method of production.

15 In addition, techniques developed for the production of "chimeric antibodies" (Morrison et al., 1984, *Proc. Natl. Acad. Sci.*, 81:6851-6855; Neuberger et al., 1984, *Nature*, 312:604-608; Takeda et al., 1985, *Nature*, 314:452-454) by splicing the genes from a mouse antibody molecule of  
20 appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a  
25 human immunoglobulin constant region. Such technologies are described in U.S. Patents Nos. 6,075,181 and 5,877,397 and their respective disclosures which are herein incorporated by reference in their entirety.

Alternatively, techniques described for the production of  
30 single chain antibodies (U.S. Patent 4,946,778; Bird, 1988, *Science* 242:423-426; Huston et al., 1988, *Proc. Natl. Acad. Sci. USA* 85:5879-5883; and Ward et al., 1989, *Nature* 334:544-546) can be adapted to produce single chain antibodies against NHP gene products. Single chain antibodies are formed by  
35 linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide.

Antibody fragments which recognize specific epitopes may be generated by known techniques. For example, such fragments include, but are not limited to: the  $F(ab')_2$  fragments which can be produced by pepsin digestion of the antibody molecule  
5 and the Fab fragments which can be generated by reducing the disulfide bridges of the  $F(ab')_2$  fragments. Alternatively, Fab expression libraries may be constructed (Huse et al., 1989, Science, 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity.

10 Antibodies to a NHP can, in turn, be utilized to generate anti-idiotypic antibodies that "mimic" a given NHP, using techniques well known to those skilled in the art. (See, e.g., Greenspan & Bona, 1993, FASEB J 7(5):437-444; and Nissinoff, 1991, J. Immunol. 147(8):2429-2438). For example  
15 antibodies which bind to a NHP domain and competitively inhibit the binding of NHP to its cognate receptor/ligand can be used to generate anti-idiotypes that "mimic" the NHP and, therefore, bind, activate, or neutralize a NHP, NHP receptor, or NHP ligand. Such anti-idiotypic antibodies or Fab  
20 fragments of such anti-idiotypes can be used in therapeutic regimens involving a NHP mediated pathway.

The present invention is not to be limited in scope by the specific embodiments described herein, which are intended as single illustrations of individual aspects of the  
25 invention, and functionally equivalent methods and components are within the scope of the invention. Indeed, various modifications of the invention, in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description. Such modifications are  
30 intended to fall within the scope of the appended claims. All cited publications, patents, and patent applications are herein incorporated by reference in their entirety.

WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule comprising at least 24 contiguous bases of nucleotide sequence first  
5 disclosed in the NHP polynucleotide described in SEQ ID NO: 1.
2. An isolated nucleic acid molecule comprising a nucleotide sequence that:
  - (a) encodes the amino acid sequence shown in SEQ ID  
10 NO: 2; and
  - (b) hybridizes under stringent conditions to the nucleotide sequence of SEQ ID NO: 1 or the complement thereof.
- 15 3. An isolated nucleic acid molecule comprising a nucleotide sequence encoding the amino acid sequence shown in SEQ ID NO:2.
- 20 4. An isolated nucleic acid molecule comprising a nucleotide sequence encoding the amino acid sequence shown in SEQ ID NO:4.

## SEQUENCE LISTING

<110> Donoho, Greg  
 Scoville, John  
 Turner, C. Alexander Jr.  
 Friedrich, Glenn  
 Zambrowicz, Brian  
 Abuin, Alejandro  
 Sands, Arthur T.

<120> Novel Human Kinase Proteins and  
 Polynucleotides Encoding the Same

<130> LEX-0103-PCT

<150> US 60/169,428  
 <151> 1999-12-07

<160> 5

<170> FastSEQ for Windows Version 4.0

<210> 1  
 <211> 1158  
 <212> DNA  
 <213> Homo sapiens

<400> 1

atggcccg	agaacggcga	gagcagctcc	tcctggaaaa	agcaagctga	agacatcaag	60
aagatcttcg	agttcaaaga	gaccctcgga	accggggcct	tttccgaagt	ggtttttagct	120
gaagagaagg	caactggcaa	gctctttgct	gtgaagtgtg	tccctaagaa	ggcgtgaag	180
ggcaaggaaa	gcagcataga	gaatgagata	gccgtcctga	gaaagattaa	gcatgaaaat	240
attgttgccc	tggagacat	ttatgaaagc	ccaaatcacc	tgtacttggt	catgcagctg	300
gtgtccgggtg	gagagctggt	tgaccggata	gtggagaagg	ggttttatac	agagaaggat	360
gccagcactc	tgatccgcca	agtcttgac	gccgtgtact	atctccacag	aatgggcatc	420
gtccacagag	acctcaagcc	cgaaaatctc	ttgtactaca	gtcaagatga	ggagtccaaa	480
ataatgatca	gtgacttttg	attgtcaaaa	atggagggca	aaggagatgt	gatgtccact	540
gcctgtggaa	ctccaggcta	tgtegtcct	gaagtcctcg	cccagaaacc	ttacagcaaa	600
gccgttgact	gctgggtccat	cggagtgatt	gcctacatct	tgctctgcgg	ctaccctcct	660
ttttatgatg	aaaatgactc	caagctcttt	gagcagatcc	tcaaggcgga	atatgagttt	720
gactctccct	actgggatga	catctccgac	tctgcaaaag	acttcattcg	gaacctgatg	780
gagaaggacc	cgaataaaaag	atacacgtgt	gagcaggcag	ctcggcaccc	atggatcgct	840
ggtgacacag	ccctcaacaa	aaacatccac	gagtcctca	gcgcccagat	ccggaaaaac	900
tttgccaaga	gcaaattggag	acaagcattt	aatgccacgg	ccgtcgtgag	acatatgaga	960
aaactacacc	tcggcagcag	cctggacagt	tcaaatgcaa	gtgtttcgag	cagcctcagt	1020
ttggccagcc	aaaaagactg	tctggcacct	tccacgctct	gtagtttcat	ttcttcttcg	1080
tcgggggtct	caggagttgg	agccgagcgg	agaccagcgc	ccaccactgt	gacggcagtg	1140
cactctggaa	gcaagtga					1158

<210> 2  
 <211> 385  
 <212> PRT  
 <213> Homo sapiens

<400> 2

Met	Ala	Arg	Glu	Asn	Gly	Glu	Ser	Ser	Ser	Ser	Trp	Lys	Lys	Gln	Ala
1				5					10					15	
Glu	Asp	Ile	Lys	Lys	Ile	Phe	Glu	Phe	Lys	Glu	Thr	Leu	Gly	Thr	Gly
			20					25					30		
Ala	Phe	Ser	Glu	Val	Val	Leu	Ala	Glu	Lys	Ala	Thr	Gly	Lys	Leu	
		35					40				45				

Phe Ala Val Lys Cys Ile Pro Lys Lys Ala Leu Lys Gly Lys Glu Ser  
 50 55 60  
 Ser Ile Glu Asn Glu Ile Ala Val Leu Arg Lys Ile Lys His Glu Asn  
 65 70 75 80  
 Ile Val Ala Leu Glu Asp Ile Tyr Glu Ser Pro Asn His Leu Tyr Leu  
 85 90 95  
 Val Met Gln Leu Val Ser Gly Gly Glu Leu Phe Asp Arg Ile Val Glu  
 100 105 110  
 Lys Gly Phe Tyr Thr Glu Lys Asp Ala Ser Thr Leu Ile Arg Gln Val  
 115 120 125  
 Leu Asp Ala Val Tyr Tyr Leu His Arg Met Gly Ile Val His Arg Asp  
 130 135 140  
 Leu Lys Pro Glu Asn Leu Leu Tyr Tyr Ser Gln Asp Glu Glu Ser Lys  
 145 150 155 160  
 Ile Met Ile Ser Asp Phe Gly Leu Ser Lys Met Glu Gly Lys Gly Asp  
 165 170 175  
 Val Met Ser Thr Ala Cys Gly Thr Pro Gly Tyr Val Ala Pro Glu Val  
 180 185 190  
 Leu Ala Gln Lys Pro Tyr Ser Lys Ala Val Asp Cys Trp Ser Ile Gly  
 195 200 205  
 Val Ile Ala Tyr Ile Leu Leu Cys Gly Tyr Pro Pro Phe Tyr Asp Glu  
 210 215 220  
 Asn Asp Ser Lys Leu Phe Glu Gln Ile Leu Lys Ala Glu Tyr Glu Phe  
 225 230 235 240  
 Asp Ser Pro Tyr Trp Asp Asp Ile Ser Asp Ser Ala Lys Asp Phe Ile  
 245 250 255  
 Arg Asn Leu Met Glu Lys Asp Pro Asn Lys Arg Tyr Thr Cys Glu Gln  
 260 265 270  
 Ala Ala Arg His Pro Trp Ile Ala Gly Asp Thr Ala Leu Asn Lys Asn  
 275 280 285  
 Ile His Glu Ser Val Ser Ala Gln Ile Arg Lys Asn Phe Ala Lys Ser  
 290 295 300  
 Lys Trp Arg Gln Ala Phe Asn Ala Thr Ala Val Val Arg His Met Arg  
 305 310 315 320  
 Lys Leu His Leu Gly Ser Ser Leu Asp Ser Ser Asn Ala Ser Val Ser  
 325 330 335  
 Ser Ser Leu Ser Leu Ala Ser Gln Lys Asp Cys Leu Ala Pro Ser Thr  
 340 345 350  
 Leu Cys Ser Phe Ile Ser Ser Ser Ser Gly Val Ser Gly Val Gly Ala  
 355 360 365  
 Glu Arg Arg Pro Arg Pro Thr Thr Val Thr Ala Val His Ser Gly Ser  
 370 375 380  
 Lys  
 385

&lt;210&gt; 3

&lt;211&gt; 1074

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 3

atggccccggg	agaacggcga	gagcagctcc	tcctggaaaa	agcaagctga	agacatcaag	60
aagatcttcg	agttcaaaga	gaccctcgga	accggggcct	tttccgaagt	ggttttagct	120
gaagagaagg	caactggcaa	gctctttgct	gtgaagtgtg	tccctaagaa	ggcgctgaag	180
ggcaaggaaa	gcagcataga	gaatgagata	gccgtcctga	gaaagattaa	gcatgaaaa	240
attgttgccc	tggaagacat	ttatgaaagc	ccaaatcacc	tgtacttggt	catgcagctg	300
gtgtccggtg	gagagctgtt	tgaccggata	gtggagaagg	ggttttatac	agagaaggat	360
gccagcactc	tgatccgcca	agtcttgga	gccgtgtact	atctccacag	aatgggcatc	420
gtccacagag	acctcaagcc	cgaaaaatctc	ttgtactaca	gtcaagatga	ggagtccaaa	480
ataatgatca	gtgactttgg	attgtcaaaa	atggagggca	aaggagatgt	gatgtccact	540
gcctgtggaa	ctccaggcta	tgctgctcct	gaagtcctcg	cccagaaacc	ttacagcaaa	600
gccgttgact	gctggtccat	cggagtgatt	gcctacatct	tgctctgcgg	ctaccctcct	660
ttttatgatg	aaaatgactc	caagctcttt	gagcagatcc	tcaaggcgga	atatgagttt	720

```

gactctccct actgggatga catctccgac tctgcaaaag acttcattcg gaacctgatg 780
gagaaggacc cgaataaaaag atacacgtgt gagcaggcag ctgggcaccc atggatcgct 840
ggtgacacag ccctcaacaa aaacatccac gagtccgtca gcgccagat ccgaaaaaac 900
tttgccaaga gcaaattggag acaagcattt aatgccacgg ccgtcgtgag acatatgaga 960
aaactacacc tcggcagcag cctggacagt tcaaatgcaa gtgtttcgag cagcctcagt 1020
ttggccagcc aaaaagactg tgcgtatgta gcaaaaccag aatccctcag ctga 1074

```

&lt;210&gt; 4

&lt;211&gt; 356

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 4

```

Met Ala Arg Glu Asn Gly Glu Ser Ser Ser Ser Trp Lys Lys Gln Ala
1      5      10      15
Glu Asp Ile Lys Lys Ile Phe Glu Phe Lys Glu Thr Leu Gly Thr Gly
20     25     30
Ala Phe Ser Glu Val Val Leu Ala Glu Glu Lys Ala Thr Gly Lys Leu
35     40     45
Phe Ala Val Lys Cys Ile Pro Lys Lys Ala Leu Lys Gly Lys Glu Ser
50     55     60
Ser Ile Glu Asn Glu Ile Ala Val Leu Arg Lys Ile Lys His Glu Asn
65     70     75     80
Ile Val Ala Leu Glu Asp Ile Tyr Glu Ser Pro Asn His Leu Tyr Leu
85     90     95
Val Met Gln Leu Val Ser Gly Gly Glu Leu Phe Asp Arg Ile Val Glu
100    105    110
Lys Gly Phe Tyr Thr Glu Lys Asp Ala Ser Thr Leu Ile Arg Gln Val
115    120    125
Leu Asp Ala Val Tyr Tyr Leu His Arg Met Gly Ile Val His Arg Asp
130    135    140
Leu Lys Pro Glu Asn Leu Leu Tyr Tyr Ser Gln Asp Glu Glu Ser Lys
145    150    155    160
Ile Met Ile Ser Asp Phe Gly Leu Ser Lys Met Glu Gly Lys Gly Asp
165    170    175
Val Met Ser Thr Ala Cys Gly Thr Pro Gly Tyr Val Ala Pro Glu Val
180    185    190
Leu Ala Gln Lys Pro Tyr Ser Lys Ala Val Asp Cys Trp Ser Ile Gly
195    200    205
Val Ile Ala Tyr Ile Leu Leu Cys Gly Tyr Pro Pro Phe Tyr Asp Glu
210    215    220
Asn Asp Ser Lys Leu Phe Glu Gln Ile Leu Lys Ala Glu Tyr Glu Phe
225    230    235    240
Asp Ser Pro Tyr Trp Asp Asp Ile Ser Asp Ser Ala Lys Asp Phe Ile
245    250    255
Arg Asn Leu Met Glu Lys Asp Pro Asn Lys Arg Tyr Thr Cys Glu Gln
260    265    270
Ala Ala Arg His Pro Trp Ile Ala Asp Thr Ala Leu Asn Lys Asn Ile
275    280    285
His Glu Ser Val Ser Ala Gln Ile Arg Lys Asn Phe Ala Lys Ser Lys
290    295    300
Trp Arg Gln Ala Phe Asn Ala Thr Ala Val Val Arg His Met Arg Lys
305    310    315    320
Leu His Leu Gly Ser Ser Leu Asp Ser Ser Asn Ala Ser Val Ser Ser
325    330    335
Ser Leu Ser Leu Ala Ser Gln Lys Asp Cys Ala Tyr Val Ala Lys Pro
340    345    350
Glu Ser Leu Ser
355

```

&lt;210&gt; 5

&lt;211&gt; 1671

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 5

gcgaaagagg	aaactgcaga	ggaggaagct	gcgccgcagc	ccgagccgcc	cggcatcccc	60
gccgcctctg	cgcccgcgcc	gcgccccggg	cgccccctcc	ccagcgcgcc	cccggccgct	120
cctccgcgcc	gggaagttgc	gccgcagccc	gagccgctcg	tcggccatgg	cccgggagaa	180
cggcgagagc	agctcctcct	ggaaaaagca	agctgaagac	atcaagaaga	tcttcgagtt	240
caaagagacc	ctcggaaccg	gggccttttc	cgaagtgggt	ttagctgaag	agaaggcaac	300
tggcaagctc	tttgctgtga	agtgtatccc	taagaaggcg	ctgaagggca	aggaaagcag	360
catagagaat	gagatagccg	tcctgagaaa	gattaagcat	gaaaaatattg	ttgccttgga	420
agacatttat	gaaagcccaa	atcacctgta	cttggtcatg	cagctgggtg	ccggtggaga	480
gctgtttgac	cggatagtgg	agaagggggt	ttatacacag	aaggatgccca	gcactctgat	540
ccgccaagtc	ttggacgcgg	tgtactatct	ccacagaatg	ggcatcgctc	acagagacct	600
caagcccgaa	aatctcttgt	actacagtca	agatgaggag	tccaaaataa	tgatcagtga	660
ctttggattg	tcaaaaatgg	agggcaaaag	agatgtgatg	tccactgcct	gtggaactcc	720
aggctatgtc	gctcctgaag	tcctcgccca	gaaaccttac	agcaaagccg	ttgactgctg	780
gtccatcgga	gtgattgcct	acatcttgct	ctgcggctac	cctccttttt	atgatgaaaa	840
tgactccaag	ctctttgagc	agatcctcaa	ggcggaatat	gagtttgact	ctccctactg	900
ggatgacatc	tccgactctg	caaaagactt	cattcggaac	ctgatggaga	aggacccgaa	960
taaaagatac	acgtgtgagc	aggcagctcg	gcacccatgg	atcgctgggtg	acacagccct	1020
caacaaaaac	atccacgagt	ccgtcagcgc	ccagatccgg	aaaaactttg	ccaagagcaa	1080
atggagacaa	gcattttaatg	ccacggccgt	cgtgagacat	atgagaaaac	tacacctcgg	1140
cagcagcctg	gacagttcaa	atgcaagtgt	ttcgagcagc	ctcagtttgg	ccagccaaaa	1200
agactgtgcg	tatgtagcaa	aaccagaatc	cctcagctga	cactgaagac	gagcctgggg	1260
tggagaggag	ggagccggca	tctgccgagc	acctcctggt	tgccaggcgc	tttctatact	1320
taatcccatg	tcatgcgacc	ctaggacttt	ttttaacatg	taatcactgg	gctgggtgca	1380
gtggctcacg	cctgtaatcc	caacactttg	ggaggctgag	gcaggaggac	tgtttgagtt	1440
caggagtttt	aagaccagcc	tgaccaacat	ggtgaaaccc	catctctact	aaaatataaa	1500
aattagccgg	gtgtgggtggc	gagcacctgt	aatgtcagct	acttgggagg	ctgaggcagg	1560
agaatcgctt	gagcccgga	gcggagggtg	caatgagctg	agatcacacc	actgcacttc	1620
agcctgggtg	acagattgag	actcctctca	aaaaaaaaaa	aaaaaaaaaa	a	1671

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
14 June 2001 (14.06.2001)

PCT

(10) International Publication Number  
**WO 01/42435 A3**

- (51) International Patent Classification<sup>7</sup>: C07K 14/00
- (21) International Application Number: PCT/US00/33240
- (22) International Filing Date: 7 December 2000 (07.12.2000)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
60/169,428 7 December 1999 (07.12.1999) US
- (71) Applicant: LEXICON GENETICS INCORPORATED  
[US/US]; 4000 Research Forest Drive, The Woodlands, TX  
77381 (US).
- (72) Inventors: DONOHO, Gregory; 41 Fallenstone Drive,  
The Woodlands, TX 77381 (US). SCOVILLE, John;  
8222 Kingsbrook Dr. #543, Houston, TX 77024 (US).  
TURNER, C., Alexander, Jr.; 67 Winter Wheat Place,  
The Woodlands, TX 77381 (US). FRIEDRICH, Glenn;  
Breland & Breland, 2207 Hermann Drive, Houston, TX  
77004 (US). ZAMBROWICZ, Brian; 18 Firethorne  
Place, The Woodlands, TX 77382 (US). ABUIN, Ale-  
jandro; 19 Belcarra Place, The Woodlands, TX 77382  
(US). SANDS, Arthur, T.; 163 Bristol Bend Circle, The  
Woodlands, TX 77382 (US).
- (74) Agents: ISHIMOTO, Lance, K. et al.; Lexicon Genetics  
Incorporated, 4000 Research Forest Drive, The Woodlands,  
TX 77381 (US).
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,  
DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR,  
HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,  
LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,  
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,  
TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian  
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European  
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,  
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,  
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- Published:  
— with international search report
- (88) Date of publication of the international search report:  
8 November 2001
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

WO 01/42435 A3

(54) Title: NOVEL HUMAN KINASE PROTEINS AND POLYNUCLEOTIDES ENCODING THE SAME

(57) Abstract: Novel human polynucleotide and polypeptide sequences are disclosed that can be used in therapeutic, diagnostic, and pharmacogenomic applications.



# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 00/33240

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> IPC 7 C07K14/00		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) IPC 7 C07K		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EMBL, BIOSIS, CHEM ABS Data		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	DATABASE EMBL 'Online! accession no. P05790, 1 November 1988 (1988-11-01) ZHOU ET AL.: "Bombyx mori fibroin heavy chain precursor." XP002167584 abstract	1-4
Y	DATABASE EMBL 'Online! accession no. 053559, 1 June 1998 (1998-06-01) COLE ET AL.: "Mycobacterium tuberculosis PGRS-family protein" XP002167585 abstract	1-4
--- -/--		
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents: *A* document defining the general state of the art which is not considered to be of particular relevance *E* earlier document but published on or after the international filing date *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) *O* document referring to an oral disclosure, use, exhibition or other means *P* document published prior to the international filing date but later than the priority date claimed *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. *&* document member of the same patent family		
Date of the actual completion of the international search 18 May 2001		Date of mailing of the international search report 05/06/2001
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Authorized officer Hardon, E

1

# INTERNATIONAL SEARCH REPORT

Int. Application No

PCT/US 00/33240

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	<p>DATABASE EMBL 'Online!  accession no. 076786,  1 November 1998 (1998-11-01)  SEZUTSU ET AL.: "Antheraea pernyi (Chinese  oak silk moth) fibroin gene. "  XP002167586  abstract</p> <p>-----</p>	1-4